18.175 Theory of Probability
Fall 2008

Section 9

Characteristic Functions. Central Limit Theorem on $\mathbb{R}$.

Let $X = (X_1, \ldots, X_k)$ be a random vector on $\mathbb{R}^k$ with distribution $\mathbb{P}$ and let $t = (t_1, \ldots, t_k) \in \mathbb{R}^k$. *Characteristic function* of X is defined by

$$f(t) = \mathbb{E}e^{i(t,X)} = \int e^{i(t,x)} d\mathbb{P}(x).$$

If X has standard normal distribution $\mathcal{N}(0,1)$ and $\lambda \in \mathbb{R}$ then

$$\mathbb{E}e^{\lambda X} = \frac{1}{\sqrt{2\pi}} \int e^{\lambda x - \frac{x^2}{2}} dx = e^{\frac{\lambda^2}{2}} \int \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\lambda)^2}{2}} dx = e^{\frac{\lambda^2}{2}}.$$

For complex $\lambda = it$, consider analytic function

$$\varphi(x) = e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \text{for} \quad x \in \mathbb{C}.$$

By Cauchy's theorem, integral over a closed path is equal to 0. Let us take a closed path $x + i0$ for x from $-\infty$ to $+\infty$ and $x + it$ for x from $+\infty$ to $-\infty$. Then

$$\begin{aligned} f(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx - \frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{it(it+x) - \frac{1}{2}(it+x)^2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2 + itx + \frac{1}{2}t^2 - itx - \frac{1}{2}x^2} dx = e^{-\frac{t^2}{2}} \int \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = e^{-\frac{t^2}{2}}. \end{aligned} \tag{9.0.1}$$

If Y has normal distribution $\mathcal{N}(m, \sigma^2)$ then

$$\mathbb{E}e^{itY} = \mathbb{E}e^{it(m+\sigma X)} = e^{itm - \frac{t^2\sigma^2}{2}}.$$

Lemma 16 *If X is a real-valued r.v. such that $\mathbb{E}|X|^r < \infty$ for integer r then $f(t) \in C^r(\mathbb{R})$ and*

$$f^{(j)}(t) = \mathbb{E}(iX)^j e^{itX}$$

for $j \leq r$.

Proof. If $r = 0$, then $|e^{itX}| \leq 1$ implies

$$f(t) = \mathbb{E}e^{itX} \to \mathbb{E}e^{isX} = f(s) \text{ if } t \to s,$$

by dominated convergence theorem. This means that $f \in C(\mathbb{R})$. If $r = 1$, $\mathbb{E}|X| < \infty$, we can use

$$\left|\frac{e^{itX} - e^{isX}}{t-s}\right| \le |X|$$

and, therefore, by dominated convergence theorem,

$$f'(t) = \lim_{s \to t} \mathbb{E}\frac{e^{itX} - e^{isX}}{t-s} = \mathbb{E}iXe^{itX}.$$

Also, by dominated convergence theorem, $\mathbb{E}iXe^{itX} \in C(\mathbb{R})$, which means that $f \in C^1(\mathbb{R})$. We proceed by induction. Suppose that we proved that

$$f^{(j)}(t) = \mathbb{E}(iX)^j e^{itX}$$

and that $r = j+1$, $\mathbb{E}|X|^{j+1} < \infty$. Then, we can use that

$$\left|\frac{(iX)^j e^{itX} - (iX)^j e^{isX}}{t-s}\right| \le |X|^{j+1}$$

so that by dominated convergence theorem $f^{(j+1)}(t) = \mathbb{E}(iX)^{j+1}e^{itX} \in C(\mathbb{R})$.

□

The main goal of this section is to prove one of the most famous results in Probability Theory.

Theorem 22 *(Central Limit Theorem) Consider an i.i.d. sequence $(X_i)_{i \ge 1}$ such that $\mathbb{E}X_1 = 0, \mathbb{E}X_1^2 = \sigma^2 < \infty$ and let $S_n = \sum_{i \le n} X_i$. Then $S_n/\sqrt{n}$ converges in distribution to $\mathcal{N}(0, \sigma^2)$.*

We will start with the following.

Lemma 17 *We have,*

$$\lim_{n \to \infty} \mathbb{E}e^{it\frac{S_n}{\sqrt{n}}} = e^{-\frac{1}{2}\sigma^2 t^2}.$$

Proof. By independence,

$$\mathbb{E}e^{it\frac{S_n}{\sqrt{n}}} = \prod_{i \le n} \mathbb{E}e^{\frac{itX_i}{\sqrt{n}}} = \left(\mathbb{E}e^{\frac{itX_1}{\sqrt{n}}}\right)^n.$$

Since $\mathbb{E}X_1^2 < \infty$ previous lemma implies that $\varphi(t) \in C^2(\mathbb{R})$ and, therefore,

$$\varphi(t) = \mathbb{E}e^{itX_1} = \varphi(0) + \varphi'(0)t + \frac{1}{2}\varphi''(0)t^2 + o(t^2) \quad \text{as} \quad t \to 0.$$

Since

$$\varphi(0) = 1,\ \varphi'(0) = \mathbb{E}iXe^{i \cdot 0 \cdot X} = i\mathbb{E}X = 0,\ \varphi''(0) = \mathbb{E}(iX)^2 = -\mathbb{E}X^2 = -\sigma^2$$

we get

$$\varphi(t) = 1 - \frac{\sigma^2 t^2}{2} + o(t^2).$$

Finally,

$$\mathbb{E}e^{\frac{itS_n}{\sqrt{n}}} = \left(\varphi\left(\frac{t}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{\sigma^2 t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n \to e^{-\frac{1}{2}\sigma^2 t^2},\quad n \to \infty.$$

□

Next, we want to show that characteristic function uniquely determines the distribution. Let $X \sim \mathbb{P}, Y \sim \mathbb{Q}$ be two independent random vectors on $\mathbb{R}^k$. We denote by $\mathbb{P} * \mathbb{Q}$ the *convolution* of $\mathbb{P}$ and $\mathbb{Q}$ which is the law $\mathcal{L}(X+Y)$ of the sum $X+Y$. We have,

$$\begin{aligned}\mathbb{P} * \mathbb{Q}(A) = \mathbb{E}I(X+Y \in A) &= \iint \mathrm{I}(x+y \in A)d\mathbb{P}(x)d\mathbb{Q}(y) \\ &= \iint \mathrm{I}(x \in A - y)d\mathbb{P}(x)d\mathbb{Q}(y) = \int \mathbb{P}(A-y)d\mathbb{Q}(y).\end{aligned}$$

If $\mathbb{P}$ has density p then

$$\begin{aligned}\mathbb{P} * \mathbb{Q}(A) &= \iint \mathrm{I}(x+y \in A) p(x) dx d\mathbb{Q}(y) = \iint \mathrm{I}(z \in A) p(z-y) dz d\mathbb{Q}(y) \\ &= \iint_A p(z-y) dz d\mathbb{Q}(y) = \int_A \Bigl(\int p(z-y) d\mathbb{Q}(y)\Bigr) dz\end{aligned}$$

which means that $\mathbb{P} * \mathbb{Q}$ has density

$$f(x) = \int p(x-y) d\mathbb{Q}(y). \tag{9.0.2}$$

If, in addition, $\mathbb{Q}$ has density q then

$$f(x) = \int p(x-y) q(y) dy.$$

Denote by $\mathcal{N}(0, \sigma^2 I)$ the law of the random vector $X = (X_1, \ldots, X_k)$ of i.i.d. $\mathcal{N}(0, \sigma^2)$ random variables whose density on $\mathbb{R}^k$ is

$$\prod_{i=1}^k \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2} x_i^2} = \Bigl(\frac{1}{\sqrt{2\pi}\sigma}\Bigr)^k e^{-\frac{1}{2\sigma^2}|x|^2}.$$

For a distribution $\mathbb{P}$ denote $\mathbb{P}^\sigma = \mathbb{P} * \mathcal{N}(0, \sigma^2 I)$.

Lemma 18 $\mathbb{P}^\sigma = \mathbb{P} * \mathcal{N}(0, \sigma^2 I)$ *has density*

$$p^\sigma(x) = \Bigl(\frac{1}{2\pi}\Bigr)^k \int f(t) e^{-i(t,x) - \frac{\sigma^2}{2}|t|^2} dt$$

where $f(t) = \int e^{i(t,x)} d\mathbb{P}(x)$.

Proof. By (9.0.2), $\mathbb{P} * \mathcal{N}(0, \sigma^2 I)$ has density

$$p^\sigma(x) = \Bigl(\frac{1}{\sqrt{2\pi}\sigma}\Bigr)^k \int e^{-\frac{1}{2\sigma^2}|x-y|^2} d\mathbb{P}(y).$$

Using (9.0.1), we can write

$$e^{-\frac{1}{2\sigma^2}(x_i - y_i)^2} = \frac{1}{\sqrt{2\pi}} \int e^{-i\frac{1}{\sigma}(x_i - y_i) z_i} e^{-\frac{1}{2} z_i^2} dz_i$$

and taking a product over $i \le k$ we get

$$e^{-\frac{1}{2\sigma^2}|x-y|^2} = \Bigl(\frac{1}{\sqrt{2\pi}}\Bigr)^k \int e^{-i\frac{1}{\sigma}(x-y,z)} e^{-\frac{1}{2}|z|^2} dz.$$

Then we can continue

$$\begin{aligned}p^\sigma(x) &= \Bigl(\frac{1}{2\pi\sigma}\Bigr)^k \iint e^{-i\frac{1}{\sigma}(x-y,z) - \frac{1}{2}|z|^2} dz d\mathbb{P}(y) \\ &= \Bigl(\frac{1}{2\pi\sigma}\Bigr)^k \iint e^{-i\frac{1}{\sigma}(x-y,z) - \frac{1}{2}|z|^2} d\mathbb{P}(y) dz \\ &= \Bigl(\frac{1}{2\pi\sigma}\Bigr)^k \int f\Bigl(\frac{z}{\sigma}\Bigr) e^{-i\frac{1}{\sigma}(x,z) - \frac{1}{2}|z|^2} dz.\end{aligned}$$

Let $z = t\sigma$.

$\square$

Theorem 23 *(Uniqueness) If*

$$\int e^{i(t,x)} d\mathbb{P}(x) = \int e^{i(t,x)} d\mathbb{Q}(x)$$

then $\mathbb{P} = \mathbb{Q}$.

Proof. By the above Lemma, $\mathbb{P}^\sigma = \mathbb{Q}^\sigma$. If $X \sim \mathbb{P}$ and $\mu \sim N(0, I)$ then $X + \sigma\mu \to X$ almost surely as $\sigma \to 0$ and, therefore, $\mathbb{P}^\sigma \to \mathbb{P}$ weakly. Similarly, $\mathbb{Q}^\sigma \to \mathbb{Q}$.

□

We proved that the characteristic function of $S_n/\sqrt{n}$ converges to the c.f. of $\mathcal{N}(0, \sigma^2)$. Also, the sequence

$$\Big(\mathcal{L}\Big(\frac{S_n}{\sqrt{n}}\Big)\Big)_{n \geq 1} \text{ - is uniformly tight,}$$

since by Chebyshev's inequality

$$\mathbb{P}\Big(\Big|\frac{S_n}{\sqrt{n}}\Big| > M\Big) \leq \frac{\sigma^2}{M^2} < \varepsilon$$

for large enough M. To finish the proof of the CLT on the real line we apply the following.

Lemma 19 *If* $(\mathbb{P}_n)$ *is uniformly tight and*

$$f_n(t) = \int e^{itx} d\mathbb{P}_n(x) \to f(t)$$

then $\mathbb{P}_n \to \mathbb{P}$ *and* $f(t) = \int e^{itx} d\mathbb{P}(x)$.

Proof. For any sequence $(n(k))$, by Selection Theorem, there exists a subsequence $(n(k(r)))$ such that $\mathbb{P}_{n(k(r))}$ converges weakly to some distribution $\mathbb{P}$. Since $e^{i(t,x)}$ is bounded and continuous,

$$\int e^{i(t,x)} d\mathbb{P}_{n(k(r))} \to \int e^{i(t,x)} d\mathbb{P}(x)$$

as $r \to \infty$ and, therefore, f is a c.f. of $\mathbb{P}$. By uniqueness theorem, distribution $\mathbb{P}$ does not depend on the sequence $(n(k))$. By Lemma 13, $\mathbb{P}_n \to \mathbb{P}$ weakly.

□